The Glasgow Aneurysm Score does not predict mortality after open abdominal aortic aneurysm in the era of endovascular aneurysm repair

Benjamin Oliver Patterson, BSc, MRCS, Alan Karthikesalingam, MA, MRCS, Robert J. Hinchliffe, MD, MRCS, Ian M. Loftus, MD, FRCS, Matt M. Thompson, MD, FRCS, and **Peter J. E. Holt, PhD, FRCS,** *London, United Kingdom*

Objective: **Endovascular aneurysm repair (EVAR) has reduced early adverse outcomes from abdominal aortic aneurysm (AAA) repair. Preferential use of EVAR may have altered the profile of patients who undergo open repair. The validity of scoring systems such as the Glasgow Aneurysm Score (GAS), devised when open surgery was the only treatment, required reappraisal.**

Methods: **Patients were identified from a database of patients undergoing elective infrarenal aneurysm repair at seven United Kingdom centers, and the GAS was calculated for each patient. Discrimination and calibration were calculated to determine the performance of the model in this setting using the C statistic, tertile analysis, and the χ^2 test. Univariate analysis was performed to determine if a new iteration of the GAS could be produced.**

Results: **We identified 330 patients who met the inclusion criteria. There were 18 deaths ≤30 days of surgery (5.4%). The average (standard deviation) GAS was 78.6 (8.8) for the survivors and 81.9 (10.4) for nonsurvivors ($P = .122$). The C statistic was 0.625 (95% confidence interval, 0.481-0.769; $P = .75$) suggesting a discriminatory ability not much better than chance alone. Despite this, calibration of the model was good. There was no significant difference in the comorbidities of either group, so no recalibration of the GAS could be performed.**

Conclusion: **The GAS did not discriminate between survivors and nonsurvivors after open AAA repair in this cohort. In the era of EVAR, it is possible that the GAS does not predict the outcome of open AAA repair. An alternative explanation is that patients with risk factors for poor outcomes from EVAR, such as adverse AAA morphology, are being selected out for open repair. (J Vasc Surg 2011;54:353-7.)**

Mortality after the rupture of abdominal aortic aneurysms (AAAs) has remained unchanged over several decades.[1-3] Early detection and elective repair is the most effective way of reducing aneurysm-related death at present. The risk associated with such procedures must be carefully considered if the procedure is to be in the best interest of the patient.

A recent systematic review of risk prediction systems for AAA repair suggested that although some systems showed promise, none of the available methods were ideal, and most had significant drawbacks.[4] The most useful score for predicting the results of open AAA repair was the Glasgow Aneurysm Score (GAS). This was first devised in 1994 and was validated using a number of patient cohorts, with varying degrees of success.

Endovascular aneurysm repair (EVAR) has reduced the early mortality and morbidity associated with AAA repair. The widespread availability, and in some centers preferential use, of EVAR may have altered the profile of patients who undergo open repair due to unsuitability for EVAR. Thus, in the endovascular era, the validity of scoring systems such as the GAS, devised for all patients with operable AAA before the development of EVAR, requires reappraisal.

The aim of this study was to quantify the discriminative ability of the GAS using contemporary multicenter UK data and attempt to recalibrate the weighting of the various components.

METHODS

Patients. Patients were identified from a database compiled for a parallel project using data from seven vascular centers in the UK encompassing the period 2005 to 2007. The six centers were St. George's Vascular Institute (London), St. Thomas' Hospital (London), Frimley Park Hospital, Hull Royal Infirmary, Queen's Medical Center (Nottingham), Bristol Royal Infirmary, and Southampton General Hospital. Institutional review board permission was previously obtained for a broader project of which this work was a component.

The inclusion criterion was elective open repair of infrarenal AAAs. Exclusion criteria included emergency admissions (mainly ruptured aneurysms), thoracoabdominal aneurysms, hybrid revascularization and aneurysm repair procedures, and any other non-AAA repair. For the purposes of these analyses, mortality was defined as death ≤30 days of the primary procedure.

From St. George's Vascular Institute.

Competition of interest: none.

Reprint requests: Benjamin Patterson, St. George's, Vascular Institute, 4th Fl, St. James' Wing, Blackshaw Rd, Tooting SW17 0QT, London, UK (e-mail: bop@doctors.net.uk).

The editors and reviewers of this article have no relevant financial relationships to disclose per the JVS policy that requires reviewers to decline review of any manuscript for which they may have a competition of interest.

0741-5214/$36.00

doi:10.1016/j.jvs.2011.01.029

The GAS was calculated for each patient identified from the database. The values for each component of the score were as described in the original study.[5] The GAS comprised patient age, + 7 for cardiovascular disease, + 10 for cerebrovascular disease, and + 14 for renal disease. We defined renal failure as an estimated glomerular filtration rate (eGRF) of <34 mL/min/1.73 m,^{2} which equates to a creatinine level >180 mg/dL. The eGFR was calculated using the Modification of Diet in Renal Disease Study Group equation.[6] Cardiovascular disease was defined as a history of angina, myocardial infarction, percutaneous coronary intervention, cardiac surgery or an abnormal electrocardiogram. Cerebrovascular disease was defined as any history of a stroke or transient ischemic attack.

Statistical analysis. Patients undergoing elective open infrarenal aneurysm repair were identified by interrogation of the database. A spreadsheet was constructed using Excel 2003 (Microsoft Corpmond, Wash). The GAS was calculated from the component variables using a formula written into the spreadsheet. All further statistical analysis was performed using SPSS 16 software (SPSS Inc, Chicago, Ill).

The measure of discrimination used was the area under a receiver-operator characteristic (ROC) curve (referred to here as the "C statistic").[6] Discrimination is defined as the ability of a test to identify high risk in individual patients. A ROC curve uses a visual representation of the sensitivity-specificity balance to determine the ability of a test to predict the outcome of a dichotomous event. Sensitivity is plotted against (1 – specificity) over the range of values for the test being investigated. An ideal test would produce a curve that would be nearly vertical, turning sharply to the horizontal near to point (0,1) with a C statistic close to 1. A diagonal line from point (0,0) to (1,1) suggests the test has a predictive power equivalent to chance alone with a C statistic of 0.5. The C statistic derived from the curve produces a value between 0.5 and 1. This represents the probability that the test will predict an outcome in comparison with random chance. Values <0.70 denote poor discriminatory power, and values >0.80 indicate reliable accuracy. A statistical test of significance produces a *P* value associated with the C statistic. This tests the null hypothesis that the C statistic will equal 0.5 and that the GAS is nondiscriminatory.

The point of the ROC closest to the upper left of the axis represents the point where the score achieves the best balance of sensitivity and specificity. It is at this point that a cutoff point can be defined. For example, when considering scores to predict the result of AAA repair, this would be the value of the score that divides patients into high-risk and low-risk subgroups.

Two methods were used to control for missing values. For categoric variables, such as the presence or absence of cerebrovascular disease, we assumed that there was no cerebrovascular disease if these data were not present. For numeric data, such serum electrolyte levels, the mean value from all samples was entered into the spreadsheet. More complex approaches to missing data were not used, such as multiple regression or value imputation. Confidence intervals (CIs) were calculated using the nonparametric asymptotic method.

Table I. Components of the Glasgow Aneurysm Score with subsequent univariate analysis

Variable[a]	*Survivors*	*Nonsurvivors*	P
Age mean (range) y	74.4 (50-94)	77 (58-87)	.108[b]
Gender, %			.062[c]
Male	85.3	100	
Female	14.7	0	
Glasgow score, mean (SD)	78.6 (10.4)	81.9 (8.792)	.122[d]
Cardiac disease	83 (26.6)	5 (27.7)	.619[d]
Cerebrovascular disease	55 (21.7)	4 (22.2)	.416[d]
Renal failure	13 (4)	1 (5.6)	.551[c]

[a]Data are presented as number (%) unless otherwise indicated.
[b]Independent samples *t* test.
[c]Fisher exact test.
[d]χ^2 test.

Because the GAS does not produce predictions of death on an individual basis, expected deaths were calculated using the Finnvasc validation cohort study cutoff points for low (<69 = 2%), moderate (69-77 = 4%), and high (>77 = 9%) risk of death.[7] Numbers of expected deaths were calculated in each tertile group based on these percentage values. This study was selected as a reference because it is one of the largest validation cohorts, and many other studies only give one cutoff value defining "high-risk" and "low-risk" groups, which is less helpful. Observed deaths were summed for each tertile. This was an assessment of calibration of the GAS in this cohort, which can be defined as the accuracy of numerical risk predictions. The χ^2 test was used to compare the observed and expected frequency of mortality in low-, moderate-, and high-risk groups.

Univariate analysis was performed on each of the four components of the GAS to determine if they were predictive of death. Dichotomous data were analyzed using the χ^2 test or Fisher exact test where there were subgroups of fewer than five patients; parametric continuous data were analyzed using the *t* test, and nonparametric continuous data were analyzed using the Mann-Whitney *U* test. Thus, χ^2 test was used for cardiac disease and cerebrovascular disease, the Fisher exact test was used for renal function, and the *t* test was used for age.

RESULTS

The database included 1243 patients who underwent open or EVAR repair of AAAs consecutively between 2005 and 2007. We identified 405 patients who underwent open repair, of which six were excluded because they underwent thoracoabdominal aneurysm repair, three had undergone iliac aneurysm repair, three had suprarenal repair, and 48 had juxtarenal aneurysms. Mortality data were not available for 15 patients. After exclusion criteria were applied, 330

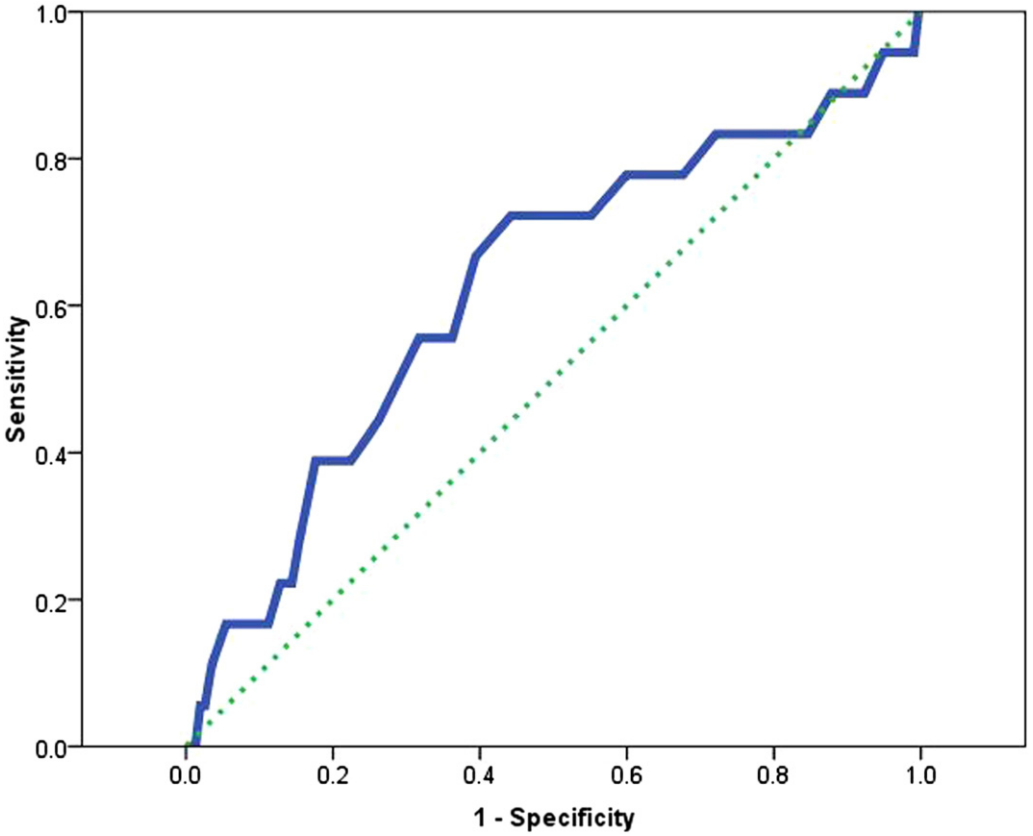

Fig 1. Receiver-operator characteristic curve displays the discriminative ability of the Glasgow Aneurysm Score to predict 30-day mortality. The *straight diagonal line* is the reference line.

Table II. Results of the area under the curve (*AUC*) calculation analysis from the receiver-operator characteristic curve shown in Fig 1[a]

		Asymptotic	
AUC	*SE*	P	*95% CI*
0.625	0.073	.75	0.481-0.769

CI, Confidence interval; *SE,* standard error.
[a]The AUC value of 0.625 suggests the Glasgow Aneurysm Score is poorly discriminatory when used to predict the 30-day mortality of patients undergoing open aneurysm repair.

patients (27% of the entire cohort) with open infrarenal AAA repair were eligible for our study. The mean age was 74.4 years and 85% were male. There were 18 (5.4%) deaths ≤30 days.

Patient characteristics are summarized dividing the population into survivors and nonsurvivors (Table I). The average (standard deviation) GAS was 78.6 (8.8) in survivors and 81.9 (10.4) in nonsurvivors. Independent sample *t* test found no significant difference between these values (95% CI −0.899 to 7.570; $P = .122$ [df 328]). An ROC analysis produced an AUC of 0.625 (95% CI, 0.481-0.769; $P = .75$). This suggested a discriminatory ability not much better than chance alone, because the *P* value was nonsignificant and the 95% CI crossed 0.5 for the AUC (Fig 1, Table II).

Observed and expected deaths were calculated for each tertile, and the significance of any discrepancy in frequency was analyzed for significance using the χ^2 test (Fig 2). Observed and expected mortality showed a good agreement, signifying that the GAS was calibrated well to the data set for low-risk patients of GAS < 69 (40 patients, 5% observed vs 2.4% expected mortality, $\chi^2 = 1.837$ [df = 1], $P = .175$), medium-risk patients with GAS 69 to 77 (111 patients, 1.8% observed vs 4.1% expected mortality, $\chi^2 = 1.451$ [df = 1], $P = .228$), and the high-risk group with GAS >77 (179 patients, 7.8% observed vs 9.3 expected, $\chi^2 = .447$ [df = 1], $P = .504$).

Univariate analysis was undertaken to determine if it would be possible to reassign the weighting of the components of the GAS using multivariate regression analysis (Table I). Neither age ($P = .108$), cardiac disease ($P = .619$), cerebrovascular disease ($P = .416$), nor renal failure ($P = .551$) were independently predictive of death. Only gender approached significance because all patients who died were male (χ^2; $P = .062$).

DISCUSSION

The GAS was developed in 1994 using a population of 500 patients undergoing AAA repair between 1980 and

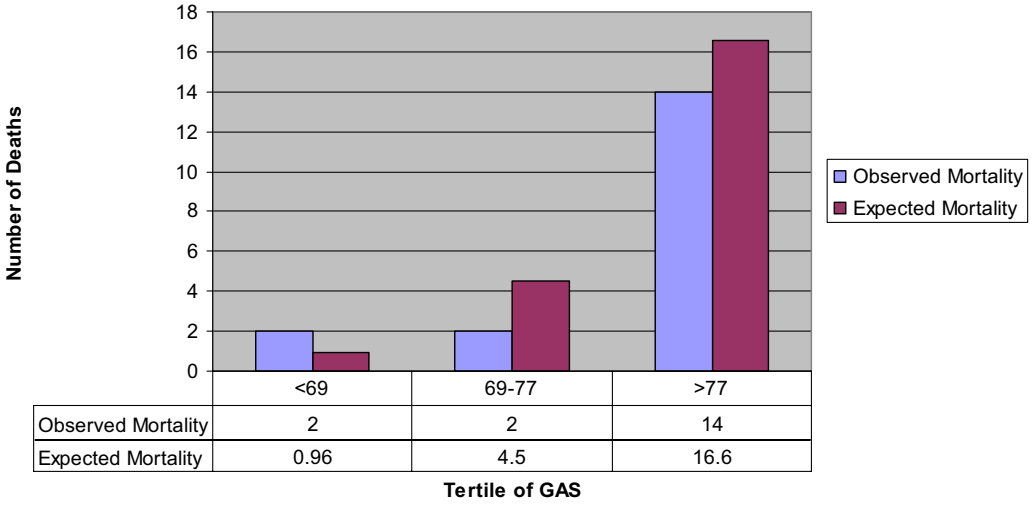

Fig 2. Observed vs expected 30-day mortality using the Finnvasc Glasgow Aneurysm Score validation tertiles as expected mortality rate. There was no significant difference between observed and expected mortality for low-risk patients of GAS <69 (n = 40), medium-risk patients with GAS 69 to 77 (n = 111), and the high-risk group with GAS >77 (n = 179).

1990. Logistic regression was used to identify factors that predicted death immediately postoperatively. Preoperative shock, myocardial dysfunction, renal impairment, and cerebrovascular disease were identified as significant predictors of death using univariate analysis and in the subsequent multivariate model.[5] The score was validated successfully by the same group shortly afterward, confirming the accuracy of this method of predicting the immediate outcome of elective and emergency open AAA repair.[8]

Further work in 2003 validated the GAS using data from 403 patients operated on in a single hospital and then again on a separate occasion using the Finnvasc national database. The AUC for immediate mortality was 0.80 (95% CI, 0.71-0.90) in the smaller group. Patient age (P = .001), cerebrovascular disease (P = .040), renal failure (P < .001), and GAS itself (P < .001) were significant predictors of death in these patients. The mean GAS was 73.1 (interquartile range [IQR], 65.2-79.6). In the survivor group, the GAS was 72.8 (IQR, 64.9-79.4) compared with the nonsurvival group mean of 86.0 (IQR, 76.4-89.6).[9]

Analysis of the Finnvasc registry demonstrated that the predictive power for mortality was lower than previously suggested with an AUC of only 0.668 (P < .0001). In this instance, cardiac disease, renal disease and patient age were significantly associated with postoperative mortality (P < .05). The average (SD) GAS was 72 (10) in the survivor group and 77 (10.4) in the nonsurvivor group. This is the only study published that firmly suggested that the GAS is inaccurate in this context.[7]

Since then, Dutch and UK investigators have shown the GAS to be highly accurate in predicting mortality (AUC >0.800).[10,11] However the Dutch study showed that none of the components of the GAS were independent predictors of mortality (all P > .05). The average GAS for this group was 73.

The GAS was a moderately accurate predictor of long-term death when applied to the open repair arm of the Dutch Randomised Endovascular Aneurysm Management (DREAM) trial (AUC = 0.740). The mean GAS was low in this group (74.7; SD, 9.3), and the proposed cutoff point between high-risk and low-risk patients was only 75.5.[12] The GAS was recently compared with more recent models and performed relatively poorly, despite an acceptable AUC of 0.749 (P = .01).[13]

The simplicity of the GAS remains a great strength, making it easier to use than any other system. It has been validated successfully previously and predicts in-hospital mortality with acceptable accuracy in historical cohorts of patients undergoing open AAA repair. A drawback of the GAS is that it does not reliably identify individual high-risk patients due to a low-positive predictive value. It is also inaccurate when used to predict morbidity.

In our current study, we have shown that the GAS may not be valid when used to predict 30-day mortality after open aneurysm repair in the era of endovascular repair. There are a number of possible explanations for this. The GAS is representative of preoperative physiologic reserve based on the burden of comorbidity present. Traditionally, this has been the mainstay of preoperative assessment for major abdominal and thoracic surgery. Factors that are not incorporated in the GAS may be important when predicting death after open repair. Composite measures of fitness, such as measurement of preoperative anaerobic threshold, may provide a more dynamic measure of preoperative physiology, but there are limited data to support this at present.[14]

It is possible that as EVAR becomes the first-line treatment for AAAs, those who undergo open repair are more likely to have complex or adverse aneurysm morphology. Some very high-risk patients with a GAS >100 survived,

which once again suggests that other factors in this group of patients not deemed suitable for EVAR may determining outcome. For example, adverse neck anatomy increases the risk of early operative reoperation in open repair of AAAs with complex neck anatomy.[15] In a study examining the late failure of open AAA repair, several factors were important, including family history of aneurysmal disease, aneurysms in other segments of the aorta, and other manifestation of atherosclerotic disease.[16] These factors may also play a part in predicting early complications. It is also possible that advances in perioperative care in the 16 years since the original GAS study mean that patients with multiple comorbidities have a greater chance of surviving surgery.

A drawback of our study is that some potentially important information was not available to us, including the reason for open repair rather than EVAR and exact cause of death in each individual was not known. Knowing this might have helped determine the reason for the higher death rate in the nonsurvival group. The reason why patients were not treated by EVAR was not known either, and there was no information regarding aneurysm morphology.

Our cohort of patients had a higher average GAS than those reported in other series, so it could that our group represents a high-risk subgroup of all patients undergoing AAA surgery. Despite this, the intraoperative mortality rate of 5.4% reported in our study was not higher than average. GAS possibly overestimated the risk of death in the survival group, or underestimated the risk of death in the nonsurvival group. The GAS was designed and validated using patients who were fit to undergo open infrarenal AAA repair in a period when this was the only treatment possible. It is feasible that it is not discriminatory in a higher-risk cohort or the specific population that are selected to undergo open repair when EVAR is an option.

CONCLUSIONS

The GAS did not discriminate between survivors and nonsurvivors after open AAA repair in this cohort. Univariate analysis of the individual score components revealed that there were no significant differences between the groups, suggesting that the score does not capture the critical predictive factors in open AAA repair. In the era of endovascular surgery, it is possible that the GAS does not predict the outcome of open AAA repair. An alternative explanation is that patients with risk factors for poor outcome from EVAR, such as adverse AAA morphology, are being selected out for open repair. Further work examining the characteristics of patients being turned down for EVAR and subsequently undergoing open repair may be justified.

We thank Professor Brian R Hopkinson, Stephen Goode (Queen's Medical Centre, Nottingham), Peter Taylor (Guy's and St. Thomas' Hospitals), David Gerard, Paul Moxey (Frimley Park Hospital), Marcus Brookes, Jude Day (Bristol Royal Infirmary), Cliff Shearman, Dee Tyler (Southampton General Hospital), Peter Gogolonceanu (St. George's Vascular Institute), Jan Polonieki (Community Health Sciences, St. George's University of London), and the BUPA Foundation.

AUTHOR CONTRIBUTIONS

Conception and design: BP, AK, PH
Analysis and interpretation: BP, AK, RH
Data collection: BP, PH
Writing the article: BP, AK, PH
Critical revision of the article: IL, MT
Final approval of the article: IL, MT
Statistical analysis: BP, AK
Obtained funding: Not applicable
Overall responsibility: BP

REFERENCES

1. Hoornweg LL, Storm-Versloot MN, Ubbink DT, Koelemay MJ, Legemate DA, Balm R. Metaanalysis on mortality of ruptured abdominal aortic aneurysms. Eur J Vasc Endovasc Surg 2008;35:558-70.
2. Ernst CB. Abdominal aortic aneurysm. N Engl J Med 1993;328:1167-72.
3. Wanhainen A, Bylund N, Bjorck M. Outcome after abdominal aortic aneurysm repair in Sweden 1994-2005. Br J Surg 2008;95:564-70.
4. Patterson BO, Holt PJ, Hinchliffe R, Loftus IM, Thompson MM. Predicting risk in elective abdominal aortic aneurysm repair: a systematic review of current evidence. Eur J Vasc Endovasc Surg 2008;36:637-45.
5. Samy AK, Murray G, MacBain G. Glasgow aneurysm score. Cardiovasc Surg 1994;2:41-4.
6. Lasko TA, Bhagwat JG, Zou KH, Ohno-Machado L. The use of receiver operating characteristic curves in biomedical informatics. J Biomed Inform 2005;38:404-15.
7. Biancari F, Heikkinen M, Lepantalo M, Salenius JP. Glasgow Aneurysm Score in patients undergoing elective open repair of abdominal aortic aneurysm: a Finnvasc study. Eur J Vasc Endovasc Surg 2003;26:612-7.
8. Samy AK, Murray G, MacBain G. Prospective evaluation of the Glasgow Aneurysm Score. J R Coll Surg Edinb 1996;41:105-7.
9. Biancari F, Leo E, Ylonen K, Vaarala MH, Rainio P, Juvonen T. Value of the Glasgow Aneurysm Score in predicting the immediate and long-term outcome after elective open repair of infrarenal abdominal aortic aneurysm. Br J Surg 2003;90:838-44.
10. Hirzalla O, Emous M, Ubbink DT, Legemate D. External validation of the Glasgow Aneurysm Score to predict outcome in elective open abdominal aortic aneurysm repair. J Vasc Surg 2006;44:712-6.
11. Tang TY, Walsh SR, Fanshawe TR, Seppi V, Sadat U, Hayes PD, et al. Comparison of risk-scoring methods in predicting the immediate outcome after elective open abdominal aortic aneurysm surgery. Eur J Vasc Endovasc Surg 2007;34:505-13.
12. Baas AF, Janssen KJ, Prinssen M, Buskens E, Blankensteijn JD. The Glasgow Aneurysm Score as a tool to predict 30-day and 2-year mortality in the patients from the Dutch Randomized Endovascular Aneurysm management trial. J Vasc Surg 2008;47:277-81.
13. Nesi F, Leo E, Biancari F, Bartolucci R, Rainio P, Satta J, et al. Preoperative risk stratification in patients undergoing elective infrarenal aortic aneurysm surgery: evaluation of five risk scoring methods. Eur J Vasc Endovasc Surg 2004;28:52-8.
14. Kothmann E, Danjoux G, Owen SJ, Parry A, Turley AJ, Batterham AM. Reliability of the anaerobic threshold in cardiopulmonary exercise testing of patients with abdominal aortic aneurysms. Anesh 2009;64:9-13.
15. Chisci E, Kristmundsson T, de Donato G, Resch T, Setacci F, Sonesson B, et al. The AAA with a challenging neck: outcome of open versus endovascular repair with standard and fenestrated stent-grafts. J Endovasc Ther 2009;16:137-46.
16. Coscas R, Greenberg RK, Mastracci TM, Eagleton M, Kang WC, Morales C, et al. Associated factors, timing, and technical aspects of late failure following open surgical aneurysm repairs. J Vasc Surg 2010;52:272-81.

Submitted Sep 28, 2010; accepted Jan 4, 2011.